Restriction endonucleases *Hind*II and *Taq*I cleave DNA with mismatched nucleotides within their recognition sequences

Josef Jiricny and Doris Martin

Friedrich Miescher Institut, PO Box 2543, CH-4002 Basel, Switzerland

Received 14 January 1986; Accepted 12 February 1986

ABSTRACT

Restriction endonucleases *Hind*II and *Taq*I, but not *Sal*I, were found to efficiently cleave synthetic hexadecanucleotide duplexes which contained either an A/C or a G/T mismatch within their respective restriction sites. Double-stranded M13 DNAs with identical mismatches were also cleaved under the assay conditions. These results suggest that the distortion of the DNA duplex, caused by these purine/pyrimidine mismatches is not sufficiently large so as to interfere with the recognition and the subsequent cleavage of the DNA by these two enzymes.

*Hind*II and *Sal*I, but not *Taq*I, were furthermore shown to hydrolyze the two strands of the duplex with different rates. The differences between the mode of recognition of their respective restriction sites by these three enzymes are discussed.

INTRODUCTION

Type II restriction endonucleases have become indispensable tools of molecular biology. Although more than 600 of these enzymes have so far been isolated and characterized (1), little is known about the precise mechanisms of their interaction with DNA; *Eco*RI, for which crystallographic data is available (2) being a notable exception. The reports which have appeared to date employed either recognition sequences containing nucleotide analogues (3) synthetic oligonucleotides (4-6) or both (7,8) in their attempts to address this complex issue. Despite these recent studies we are still largely unable to define the precise structural features of the recognition sites which are necessary for binding and phosphodiester bond cleavage. Furthermore, although most of these enzymes exist in the form of dimers (9) and recognise palindromic sequences of similar size (typically 4-6 nucleotides), they show complex diversity of interaction with these sites (3).

We are currently investigating the effects of changes introduced into a single such site, GTCGAC, on three enzymes capable of cleaving DNA within this sequence: HindII, SalI and TaqI, which recognize GTYRAC, GTCGAC and TCGA respectively (R=purine, Y=pyrimidine).

The substitution of the central guanine with adenine, and/or of its complementary cytosine with thymine results in the generation of three new duplexes, two of which contain a purine/pyrimidine mismatch within the restriction site. The sequences of the 16-mer duplexes containing these recognition sites are shown in figure 1.

MATERIALS AND METHODS

The restriction enzymes used in this study were purchased from Boehringer Mannheim and used according to the manufacturer's instructions.

Oligonucleotide synthesis: The 16-mer oligodeoxyribonucleotides were synthesized on an Applied Biosystems model 380A synthesizer and purified by polyacrylamide gel electrophoresis according to Atkinson and Smith (10).

Duplex construction: a,b The 5'-[³²P]-labelled oligonucleotide (20pmole) was annealed to its complementary 16-mer (25pmole) (80°C, 5min., then 20°C, 20 min.) in annealing buffer (10mM Tris pH 8.3, 10mM MgCl₂) in a total volume of 10ul.

c The nicked-duplex DNA containing a G/T mismatch was constructed by hybridization of a denatured BalI-restricted M13 mp9 replicative form (RF) DNA with 5 mole-equivalents of M13 mp9/op7 phage (ss) DNA (a mp9 opal mutant with a C-T transition mutation at position 6251). The A/C duplex was obtained by annealing the BalI-cut M13 mp9/op7 RF with excess M13mp9 phage DNA (11).

Enzyme digests of duplex oligonucleotides: The oligonucleotide duplexes (4pmole) were suspended in the appropriate assay buffer and digested in the presence of BSA (0.5mg/ml) with 1 unit of the appropriate restriction enzyme for up to 3 hours at 37°C in a total volume of 20ul. 2ul aliquots were withdrawn from the reaction mixture after 10, 30, and 180 minutes; the enzyme was inactivated by the addition of formamide loading dye (4ul), followed by heating at 90°C for 3 min. The progress of the reaction was

monitored by electro-phoresis on a 20% denaturing polyacrylamide gel. Following autoradiography, the radioactive bands were cut out and the relative yield of product was determined by Cerenkov counting, using the relationship

$$[\% \text{Product}] = \text{cpm}(\text{product}) / [\text{cpm}(\text{product}) + \text{cpm}(\text{substrate})].$$

The results are shown in Table 1.

Enzyme digests of nicked duplexes: The duplex DNA (0.2 pmole), suspended in reaction buffer was subjected to a digest with 1 unit of *Hind*II and 1 unit of *Bal*I at 37°C for 1 hour in a total volume of 20ul. At the end of this period the reaction was stopped by the addition of 5ul of glycerol loading dye and the samples were analyzed by electrophoresis on a 0.8% agarose gel containing 1mg/l of ethidium bromide.

RESULTS AND DISCUSSION.

All three enzymes were shown to digest the oligonucleotide duplexes containing their respective bona fide restriction sequences (Table 1, R/Y=G/C and, in the case of *Hind*II only, A/T). By selectively 5'-[³²P]-labelling either the top or the bottom strand (fig. 1a,b respectively), we were able to show that *Sal*I and *Hind*II do not cleave both the strands of the duplex with equal rates. It was reported some years ago that a number of enzymes appeared to cleave different sites within the same DNA

Table 1: *Sal*I, *Hind*II and *Taq*I digests of oligonucleotide duplexes a and b. The figures quoted in this table represent the average values obtained from three separate experiments; the observed deviations between two successive experiments were typically within +/- 8% of the quoted values.

Duplex	R/Y	Yield of product(%)								
		<i>Sal</i> I			<i>Hind</i> II			<i>Taq</i> I		
		10'	30'	180'	10'	30'	180'	10'	30'	180'
a	G/C	36	64	84	26	48	94	14	31	97
	G/T	0	0	0	4	13	61	10	21	77
	A/C	0	0	0	26	53	88	8	19	86
	A/T	0	0	0	21	41	88	0	0	0
b	G/C	13	20	74	17	26	72	15	33	95
	G/T	0	0	0	7	16	22	11	22	78
	A/C	0	0	0	16	30	56	8	20	75
	A/T	0	0	0	18	32	75	0	0	0

Figure 1: The structure of the 16-mer oligodeoxyribonucleotide duplexes (a,b), and the nicked-duplex M13 substrates (c). [p denotes a [32 P]-phosphate group, R denotes either purine, Y denotes either pyrimidine]

molecule with different rates (4,12,13), and this phenomenon has been attributed to the influence of the sequences flanking the sites. Yoo et al. (5) and others (6,13) demonstrated that the rates of DNA cleavage within a single site may vary between the two strands, with some enzymes showing a preference for purines (*MspI*), others (*HpaII*) for pyrimidines as the flanking nucleotides. Our results show that both, the *SalI* and the *HindII* upper-strand restriction sites, which are flanked on both sides by pyrimidines (fig.1), are cleaved more readily (Table 1a, R/Y=G/C) than the lower strand sequences (Table 1b, R/Y=G/C). Thus both *HindII* and *SalI* fall into the latter, pyrimidine-preferring group. By contrast, *TagI* cleaved both the oligonucleotide strands with similar kinetics (Table 1a,b, R/Y=G/C). In this case, neighbouring sequences do not seem to influence the recognition/cleavage process.

As can be seen from Table 1, *SalI* was able to cleave only its bona fide recognition sequence (R/Y=G/C). *HindII* was, on the other hand, equally capable of cleaving the duplexes which contain either the A/C or the G/T mismatch within the recognition site (Table 1, R/Y=A/C, G/T). This was an unexpected finding. However, this enzyme can accommodate either pyrimidine in the third, and either purine in the fourth position of its recognition sequence, GTRYAC, suggesting that only a single hydrogen-bonding site per central base-pair (possibly the N-7 of the purine) is required for stabilization of the protein-

DNA complex. Thus, should this enzyme act as a dimer, the recognition requirements of each subunit would be fulfilled by all four oligonucleotides (fig.1). The distortion, arising within the duplex restriction site due to the presence of the mismatch, is apparently not sufficient to interfere with the recognition and the subsequent cleavage of this site by HindII. Indeed, the structure of a G/T mismatch, recently obtained from nmr (14) and X-ray diffraction studies (15), shows that the distortion of the duplex DNA caused by this mismatch is rather small. Viewed in this light, the fact that this enzyme is capable of cleaving a purine/pyrimidine mismatch should not be altogether unexpected.

By contrast, the four-base recognition sequence of TagI (TCGA) apparently allows for no ambiguities. Surprisingly, we could show that this enzyme is, like HindII, capable of cleaving the 16-mer duplex with either the G/T or the A/C mismatch within the restriction site. In this instance, the presence of a bona fide recognition sequence in one strand of the duplex seems to be sufficient for recognition and cleavage.

These data, and those of Bodnar et al. (3) indicate that TagI recognizes and binds to a single DNA strand only. The enzyme is extremely sensitive to changes introduced into both strands of its recognition site, being able to cleave neither TCG^{Me}A (16), nor TCIA (I=inosine, J.J., unpublished results). On the other hand, changes involving only one strand of the tetramer palindrome seem to be of little importance. Bodnar et al. (3) have shown it to be virtually unaffected by a number of minor and major groove modifications introduced into one strand of its restriction site by a variety of nucleotide analogues. Nevertheless, as the enzyme is unable to hydrolyze ss DNA (J.J., unpublished), both strands must be required for cleavage of the sugar-phosphate backbone. The fact that the optimum working temperature of this enzyme is 85°C (T. Bickle, personal communication), at which DNA would be expected to possess considerable single-stranded character, lends support to the above proposal.

In order to establish that the behaviour of TagI and HindII was not restricted solely to oligonucleotide duplexes, we con-

Figure 2: HindII/BalI digests of the nicked-duplex M13 molecules c.

Lane 1: lambda/HindIII marker; lane 2: M13 mp9 RF digested with BalI; lane 3: M13 mp9 RF digested with BalI and HindII; lane 4: nd DNA R/Y=G/T [these lanes show the composition of the unpurified reaction mixture obtained upon hybridization of viral DNA (ss DNA) with BalI-cut M13 RF DNA (7229 bp) to give the nicked-duplex molecules (nd DNA)] ; lane 5: 4 digested with BalI and HindII; lane 6: nd DNA R/Y=A/C; lane 7: 6 digested with BalI and HindII.

structed two nicked-duplex (nd) molecules containing a single G/T or A/C mismatch within the M13 mp9 HindII site (fig.1c). The digest of these duplexes with HindII and BalI should produce two fragments, 6060 and 1169 base-pairs long. If the mismatch were not digested, only one fragment, 7229 bp long, would be expected. Figure 2 clearly shows that both the nicked-duplex molecules (lane 4&5 R/Y=G/T, lane 6&7 R/Y=A/C) were digested to yield the same size fragments as the control M13 RF (lane 3), indicating that the DNA was cut in the mismatched restriction sites. [TagI was not examined in this latter assay, as it cleaves RF M13 mp9 DNA at 11 different positions.]

CONCLUSIONS

HindII and TagI are presumably the first identified representatives of a much larger group of restriction enzymes, which

tolerate mismatches within their recognition sites. Indeed, approximately 10% of the 200 or so commercially available enzymes recognize multiple sequences. It is not inconceivable that these enzymes could also be capable of cleaving DNA containing mismatches within their recognition sequences. Although it is difficult to speculate on the biological significance of this phenomenon in the bacterial restriction/modification systems, the information obtained from these, and future experiments of this type, may prove invaluable in the elucidation of the mechanisms of DNA sequence recognition not only by type II restriction enzymes, but by DNA binding enzymes in general.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to K. Wiebauer for many helpful discussions, and to T. Bickle, B. Hohn and D. Monard for critical reading of the manuscript.

REFERENCES

1. Kessler, C.; Neumaier, P.S.; Wolf, W.; (1985), *Gene* **33**, 1-102.
2. Frederick, C.A.; Grable, J.; Melia, M.; Samudzi, C.; Jen-Jacobson, L.; Wang, B.-C.; Greene, P.; Boyer, H.W.; Rosenberg, J.M.; (1984) *Nature*, **309**, 327-331.
3. Bodnar, J.W.; Zempsky, W.; Warder, D.; Bergson, C.; Ward, D.C.; (1983) *J.Biol.Chem.*, **258**, 15206-15213.
4. Hofer, B.; Koster, H.; (1980), *Nucl.Acids Res.*, **8**, 6143-6162.
5. Yoo, O.J.; Agarwal, K.L.; (1980), *J.Biol.Chem.*, **255**, 10559-10562.
6. Baumstark, B.R.; Roberts, R.J.; RajBhandary, V.L.; (1979) *J.Biol.Chem.*, **254**, 8943-8950.
7. Ono, A.Y.; Ueda, T.; (1984) *Nucl.Acids Res.*, **12**, 8939-8949.
8. Connolly, B.A.; Potter, B.V.L.; Eckstein, F.; Pingoud, A.; Grotjahn, L.; (1984) *Biochemistry*, **23**, 3443-3453.
9. Modrich, P.; Roberts, R.J.; (1982) *Nucleases* (Linn & Roberts eds.), CSH Monograph series, **14**, 109-154.
10. Atkinson, T.; Smith, M.; (1984) *Oligonucleotide synthesis* (M.J. Gait ed.), practical approach series, IRL press Oxford, Washington DC, 35-81.
11. Kramer, B.; Kramer, W.; Fritz, H.-J.; (1984) *Cell*, **38**, 879-887.
12. Thomas, M.; Davis, R.W.; (1975) *J.Mol.Biol.*, **91**, 315-328.
13. Rubin, R.A.; Modrich, P.; (1978) *Nucl.Acids Res.*, **5**, 2991-2997.
14. Patel, D.J.; Kozlowski, S.A.; Marky, L.A.; Rice, J.A.; Broka, C.; Dallas, J.; Itakura, K.; Breslauer, K.J.; (1982) *Biochemistry*, **21**, 437-444.
15. Brown, T.; Kennard, O.; Kneale, G.; Rabinovich, D.; (1985) *Nature*, **315**, 604-606.
16. Horiuchi, K.; Zinder, N.D.; (1975) *Proc.Nat.Acad.Sci.USA*, **72**(7), 2555-2558.